Investigating the quality of video consultations performed using fourth generation (4G) mobile telecommunications

Liam J Caffery and Anthony C Smith
Centre for Online Health
The University of Queensland

Overview

• Background
  – What is 4G?
  – Why is the study important?
  – Packet loss and jitter
• Methods
• Results
• Conclusion

What is 4G?

• Fourth Generation (4G) Mobile Broadband
• IMT-Advanced definition
• Long Term Evolution (LTE) Advanced
• Higher throughput
• Lower latency

4G Vs 3G

<table>
<thead>
<tr>
<th></th>
<th>3G</th>
<th>4G</th>
</tr>
</thead>
<tbody>
<tr>
<td>Requirement driving architecture</td>
<td>Predominantly voice (data was an add on)</td>
<td>Converged data and voice over IP</td>
</tr>
<tr>
<td>Switching</td>
<td>Circuit (voice) and packet (data)</td>
<td>Packet only including voice</td>
</tr>
<tr>
<td>Download speeds</td>
<td>Max 384 kbps, Typical 200 kbps</td>
<td>Max 100 Mbps, Typical 20 - 70 Mbps</td>
</tr>
<tr>
<td>Upload</td>
<td>Max 384 Kbps, Typical &lt; 200 kbps</td>
<td>Max 50 Mbps, Typical 4 – 12 Mbps</td>
</tr>
<tr>
<td>Latency</td>
<td>20ms</td>
<td>10ms</td>
</tr>
</tbody>
</table>

Why is it important?

• Audio and video is more susceptible to network conditions
• NBN
  – FTTN Vs FTTP
  – green field Vs brown field
  – speed of rollout
• DSL
  – Condition of copper network
  – Continued maintenance

Coverage

• Wireless broadband uses radio waves from antennae on fixed towers.
Packet switched

- Used by Internet Protocol (IP) networks
- Problems inherent IP networks
  - Packet Loss
  - Jitter

Packet loss

- Failure of data packet to arrive at destination or data packet to arrive out of order

Packet loss

- Packet Loss affect both audio and video

Jitter

- Uneven rate of arrival of audio or video packets
Aim

- Determine if 4G is a suitable broadband technology to use for clinical video consultations
- To identify if variation in perceived audio and video quality are due to underlying network performance.

Methods

- International Telecommunications Union (ITU) scales
- Measured subjective audio and video quality of actual video consultations from perspective of the clinician

Audio scales

<table>
<thead>
<tr>
<th>Score</th>
<th>Audio quality</th>
<th>Listening effort</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>Excellent</td>
<td>Complete relaxation possible; no effort required</td>
</tr>
<tr>
<td>4</td>
<td>Good</td>
<td>Attention necessary; no appreciable effort required</td>
</tr>
<tr>
<td>3</td>
<td>Fair</td>
<td>Moderate effort required</td>
</tr>
<tr>
<td>2</td>
<td>Poor</td>
<td>Considerable effort required</td>
</tr>
<tr>
<td>1</td>
<td>Bad</td>
<td>No meaning understood with any feasible effort</td>
</tr>
</tbody>
</table>

Video scales

<table>
<thead>
<tr>
<th>Score</th>
<th>Video quality</th>
<th>Image impairment</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>Excellent</td>
<td>Imperceptible</td>
</tr>
<tr>
<td>4</td>
<td>Good</td>
<td>Perceptible, but not annoying</td>
</tr>
<tr>
<td>3</td>
<td>Fair</td>
<td>Slightly annoying</td>
</tr>
<tr>
<td>2</td>
<td>Poor</td>
<td>Annoying</td>
</tr>
<tr>
<td>1</td>
<td>Bad</td>
<td>Very annoying</td>
</tr>
</tbody>
</table>

Data collection

<table>
<thead>
<tr>
<th>Site 1 – Nudgee</th>
<th>Site 2 – Chinchilla</th>
<th>Site 3 – Townsville</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distance from Brisbane (km)</td>
<td>18</td>
<td>291</td>
</tr>
<tr>
<td>Model</td>
<td>Cisco ™ SX20</td>
<td>Cisco ™ SX20</td>
</tr>
<tr>
<td>Call protocol</td>
<td>SIP/RTSP</td>
<td>SIP/RTSP</td>
</tr>
<tr>
<td>Call rate (kbps)</td>
<td>512</td>
<td>384</td>
</tr>
<tr>
<td>Audio codec</td>
<td>AAC LD @ 84</td>
<td>AAC LD @ 84</td>
</tr>
<tr>
<td>Video codec</td>
<td>H264</td>
<td>H264</td>
</tr>
<tr>
<td>Video bitrate (kbps)</td>
<td>448</td>
<td>320</td>
</tr>
<tr>
<td>Video resolution</td>
<td>1280 x 720 @ 30fps</td>
<td>720 x 480 @ 30fps</td>
</tr>
</tbody>
</table>

Setting
Setting

<table>
<thead>
<tr>
<th>Site</th>
<th>Distance from Brisbane</th>
<th>Latency</th>
<th>Upload</th>
<th>Download</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nudgee</td>
<td>18 km</td>
<td>17 - 22 ms</td>
<td>4 - 12 Mbps</td>
<td>27 - 41 Mbps</td>
</tr>
<tr>
<td>Chinchilla</td>
<td>291 km</td>
<td>42 - 56 ms</td>
<td>3 - 5 Mbps</td>
<td>24 - 34 Mbps</td>
</tr>
<tr>
<td>Townsville</td>
<td>1358 km</td>
<td>50 - 52 ms</td>
<td>11 - 15 Mbps</td>
<td>27 - 29 Mbps</td>
</tr>
</tbody>
</table>

MOS

- Mean Opinion Score (MOS)
  - Audio
  - Video
  - Combined (audio + video)

\[
\text{MOS}_{\text{combined}} = 0.765 + 0.1925 \times \text{MOS}_{\text{audio}} \times \text{MOS}_{\text{video}} [1]
\]


Results

- N=36
- Jan – Nov 2014
- duration = 50 hours
- \text{MOS}_{\text{audio}} = 4.1 \pm 0.62
- \text{MOS}_{\text{video}} = 4.4 \pm 0.32
- \text{MOS}_{\text{combined}} = 4.2 \pm 0.62

Network metrics

- Jitter and Packet Loss
  - Audio and video
- Correlation with quality scores
Results

- Audio packet loss < 1% (max 0.65%) *
- Video packet loss < 1% (max 0.63%)
- Maximum audio jitter 14 – 57 ms
- Maximum video jitter 14 – 46 ms **

*Packet loss <1%
**Jitter < 150 ms
Conclusions

- MOS Satisfied to Very Satisfied
- Packet loss < 1%
- Jitter < 60 ms
- Listening effort
  - 97% No effort or no appreciable effort required
- Image impairment
  - 97% Imperceptible or perceptible but not annoying
- Little correlation between MOS and packet loss and jitter indicating below level

Conclusions

- 4G is a suitable broadband technology for video consultations
- Variations in perceived video and audio quality are not due to underlying network performance

Contact

Liam Caffery
l.caffery@uq.edu.au
www.uq.edu.au/coh